

Validación de una Herramienta de Evaluación Basada en el Modelo Rasch para Medir la Resolución Creativa de Problemas en Estudiantes Mediante el Uso de TIC

Rasch Measurement Validation of an Assessment Tool for Measuring Students' Creative Problem-Solving through the Use of ICT

 **Farida Farida**
Universitas Islam Negeri Raden Intan Lampung. Indonesia

 **Yosep Aspat Alamsyah**
Universitas Islam Negeri Raden Intan Lampung. Indonesia

 **Bambang Sri Anggoro**
Universitas Islam Negeri Raden Intan Lampung. Indonesia

 **Tri Andari**
Universitas PGRI Madiun. Indonesia

 **Restu Lusiana**
Universitas PGRI Madiun. Indonesia

Recibido: 2024/04/03; **Revisado:** 2024/05/14; **Aceptado:** 2024/07/10; **Online First:** 2024/07/19; **Publicado:** 2024/09/01

RESUMEN

A pesar del creciente reconocimiento de la importancia de la resolución creativa de problemas (CPS) a través del uso de las TIC en la educación con un currículo independiente, existe una falta de validación psicométrica integral para los instrumentos de evaluación de CPS. Este estudio tuvo como objetivo desarrollar y evaluar un instrumento de evaluación para medir la CPS a través del uso de las TIC en estudiantes, utilizando el modelo Rasch. Participaron un total de 137 estudiantes de educación superior como encuestados. Para este propósito, se crearon 20 ítems que cubrían diferentes aspectos de la CPS. El análisis de datos se realizó utilizando el software Winstep y SPSS. Se empleó el modelo Rasch para confirmar la validez y fiabilidad del instrumento de medición recién desarrollado. Los hallazgos del análisis del modelo Rasch indicaron un buen ajuste entre los ítems de evaluación y los estudiantes individuales. Los ítems demostraron un ajuste adecuado con el modelo Rasch, lo que permitió diferenciar los niveles de dificultad entre diferentes ítems y mostró un nivel satisfactorio de fiabilidad. El análisis del mapa de Wright reveló patrones de interacción entre los ítems y los individuos, discriminando efectivamente entre los diversos niveles de habilidades de los estudiantes. En particular, un ítem mostró DIF basado en el género, lo que favorece a los estudiantes varones en términos de sus habilidades de respuesta. Además, el estudio identificó que las estudiantes en el cuarto semestre exhibieron habilidades de respuesta promedio más altas en comparación con las estudiantes en el sexto y octavo semestre. Además, se observaron diferencias significativas en las habilidades de respuesta entre estudiantes varones y mujeres, así como entre estudiantes que residen en áreas urbanas y rurales. Estos hallazgos son cruciales para los educadores, enfatizando la necesidad de implementar estrategias de diferenciación efectivas.

ABSTRACT

Despite increasing recognition of the importance of creative problem solving (CPS) through the use of ICT in independent curriculum education, there is a lack of comprehensive psychometric validation for CPS assessment instruments. This study aimed to develop and evaluate an assessment instrument to measure CPS through the use of ICT students using the Rasch model. A total of 137 higher education students participated as respondents. For this purpose, 20 items were created, covering different aspects of CPS. Data analysis was performed using Winstep and SPSS software. The Rasch model was employed to confirm the validity and reliability of the newly developed measurement instrument. The findings of the analysis of the Rasch model indicated a good fit between the assessment items and individual students. The items demonstrated adequate fit with the Rasch model, allowing for differentiation of difficulty levels among different items and exhibiting a satisfactory level of reliability. The Wright map analysis revealed patterns of interaction between the items and individuals, effectively discriminating between varying levels of student abilities. In particular, an item showed DIF based on gender, which favours male students in terms of their response abilities. Furthermore, the study identified that female students in the fourth semester exhibited higher average response abilities compared to female students in the sixth and eighth semesters. Furthermore, significant differences in response abilities were observed between male and female students, as well as between students who resides in urban and rural areas. These findings are crucial for educators, emphasising the need to implement effective differentiation strategies.

PALABRAS CLAVES · KEYWORDS

Resolución creativa de problemas, uso de TIC, educación, género, psicometría, medición Rasch
Creative problem solving, the use of ICT, education, gender, psychometric, Rasch measurement

1. Introducción

En los últimos años, ha habido un creciente interés en evaluar las habilidades de resolución creativa de problemas (CPS) entre los estudiantes utilizando tecnologías de la información y la comunicación (TIC), ya que se reconoce como una habilidad crucial en el siglo XXI (Care & Kim, 2018; Hao et al., 2017). La capacidad de pensar creativamente y encontrar soluciones innovadoras a problemas complejos es crucial en un mundo en constante cambio que exige adaptabilidad y pensamiento creativo (Suherman & Vidákovich, 2022). Además, las TIC desempeñan un papel central en el marco DigCompEdu, donde las tecnologías se integran en las prácticas de enseñanza de manera pedagógicamente significativa (Caena & Redecker, 2019). Comprender y promover las habilidades de CPS entre los estudiantes es crucial por varias razones. En primer lugar, fomentar la creatividad capacita a las personas para generar soluciones innovadoras (Lorusso et al., 2021), fomentar el espíritu empresarial (Val et al., 2019), e impulsar el crecimiento económico (Florida, 2014). Además, la resolución creativa de problemas (CPS) es vital para abordar desafíos sociales complejos, como la sostenibilidad y la desigualdad social (Mitchell & Walinga, 2017).

Como implicación de la era del siglo XXI, numerosos países han reconocido la necesidad de incorporar habilidades como la resolución creativa de problemas (D. Lee & Lee, 2024), pensamiento computacional (Küçükaydın et al., 2024), Tecnologías de la Información y la Comunicación (TIC) (Rahimi & Oh, 2024), y creatividad (Suherman & Vidákovich, 2024), que son identificadas como habilidades esenciales para el siglo XXI. Estas competencias están siendo cada vez más integradas en los currículos educativos para preparar a los estudiantes para las demandas de la sociedad moderna y el mercado laboral en evolución (Abina et al., 2024). Por lo tanto, el énfasis en desarrollar estas habilidades refleja un reconocimiento global de su importancia para el éxito futuro (Yu & Duchin, 2024). El impacto de la era del siglo XXI se extiende más allá de la educación hacia la fuerza laboral y la vida diaria. La investigación de Arredondo-Trapero et al. (2024) enfatiza que la resolución de problemas, el pensamiento crítico y las TIC son cruciales para la innovación y la competitividad en el ámbito global de la educación. En consecuencia, los sistemas educativos y los currículos están bajo presión para reformarse y equipar a los estudiantes con estas habilidades, asegurando así que estén preparados para los desafíos futuros.

La resolución creativa de problemas (CPS) en un currículo independiente también fomenta la colaboración y el trabajo en equipo. Al incorporar CPS en el currículo, los estudiantes tienen la capacidad de abordar problemas con una mente abierta y explorar múltiples perspectivas (Burns & Norris, 2009). Se les anima a cuestionar suposiciones, desafiar el saber convencional y buscar soluciones alternativas. Este proceso no solo desarrolla habilidades analíticas, sino que también fomenta la creatividad y habilidades de pensamiento divergente (Suherman & Vidákovich, 2022). Un currículo independiente proporciona a los estudiantes la libertad de explorar temas de interés y participar en el aprendizaje autodirigido (Lestari et al., 2023), colaborativo y en equipo (Zheng et al., 2024). Este enfoque se alinea con las necesidades del estudiante del siglo XXI al enfatizar caminos educativos personalizados que responden a las fortalezas y preferencias individuales (Zhang et al., 2024). Se anima a los estudiantes a trabajar juntos, aprovechando sus diversas perspectivas y habilidades para abordar desafíos complejos (Utami & Suswanto, 2022). Este entorno colaborativo fortalece las habilidades interpersonales y de comunicación, preparándolos para futuras colaboraciones empresariales. Además,

incorporar CPS en el currículo prepara a los estudiantes para las demandas de la fuerza laboral rápidamente cambiante del siglo XXI (Stankovic et al., 2017). A medida que el mundo se vuelve cada vez más complejo e interconectado (Brunner et al., 2024), los empleadores buscan individuos que puedan pensar críticamente (Carnevale & Smith, 2013), adaptarse al cambio y generar soluciones innovadoras (Wolcott et al., 2021).

La resolución creativa de problemas (CPS) ha ganado reconocimiento como un conjunto de habilidades valioso en contextos educativos, incluyendo el currículo independiente. Sin embargo, su implementación puede enfrentar varios desafíos que deben abordarse para asegurar su efectividad y éxito. Investigaciones previas han resaltado la importancia de las habilidades de CPS en diversos contextos educativos (Greiff et al., 2013; Wang et al., 2023; Wolcott et al., 2021). Uno de los desafíos es la falta de capacitación docente y familiaridad con las técnicas de CPS. Estudios han demostrado que los educadores pueden tener dificultades para integrar CPS en sus prácticas pedagógicas debido a conocimientos limitados y poca experiencia en facilitar actividades de CPS (van Hooijdonk et al., 2020). Esto puede obstaculizar la implementación efectiva de CPS y limitar su impacto en el aprendizaje de los estudiantes. Además, la investigación ha explorado factores que contribuyen al desarrollo de las habilidades de CPS, incluyendo la influencia de la cultura (Cho & Lin, 2010), enfoques instructivos y características individuales (Samson, 2015). Sin embargo, se entienden de manera limitada los factores específicos que desarrollan las habilidades de CPS en los estudiantes.

En cuanto a la evaluación, estudios previos han explorado enfoques alternativos para evaluar las habilidades de CPS. Se han propuesto métodos de evaluación basados en el desempeño, portafolios y rúbricas que evalúan la creatividad, el pensamiento crítico, el pensamiento metafórico, la motivación tecnológica relacionada con la resolución de problemas y las habilidades para resolver problemas como métodos más completos y auténticos (Abosalem, 2016; Farida et al., 2022; Liu et al., 2024; Montgomery, 2002; Suastra et al., 2019). Estos enfoques proporcionan una visión más holística de las habilidades de CPS de los estudiantes y fomentan el desarrollo de habilidades de pensamiento de orden superior. Sin embargo, la investigación también ofrece soluciones potenciales, como el desarrollo profesional de los docentes, la integración en el currículo y métodos de evaluación alternativos. Al revisar la literatura relevante, nuestro objetivo es construir sobre el conocimiento existente e identificar las lagunas en la comprensión, proporcionando una base para la contribución de este estudio al campo.

A pesar del creciente reconocimiento de la importancia de la resolución creativa de problemas (CPS) en la educación con currículo independiente, existe una falta de validación psicométrica integral de instrumentos de evaluación de CPS. Validar el instrumento de medición es crucial, ya que CPS sigue siendo un constructo psicológico poco definido desde una perspectiva psicométrica (Tang et al., 2020). En ausencia de evaluaciones válidas y confiables, los instructores enfrentan desafíos para medir de manera confiable el aprendizaje de CPS de los estudiantes en el aula. Por lo tanto, este estudio tiene como objetivo validar CPS utilizando el modelo de Rasch, investigando si los datos se alinean con la medición del modelo de Rasch. Las preguntas de investigación son las siguientes:

1. ¿El instrumento desarrollado demuestra fiabilidad y validez basadas en la medición de Rasch?
2. ¿Cuáles son los patrones de interacción entre ítems y personas en el instrumento desarrollado según el mapa de Wright?

3. ¿Existen sesgos en el instrumento basados en el género según el análisis de Funcionamiento Diferencial del Ítem (DIF)?
4. ¿Cómo se desarrolla la resolución colaborativa de problemas (CPS) en los estudiantes en términos de calificaciones del curso?

1.1. CPS

La resolución creativa de problemas (CPS) es un proceso que permite a las personas aplicar el pensamiento creativo y crítico para encontrar soluciones a problemas cotidianos (T. Lee et al., 2023; Van Hooijdonk et al., 2023). CPS ayuda a eliminar la tendencia a abordar los problemas de manera aleatoria y, como resultado, evitar sorpresas y/o decepciones con las soluciones. Los estudiantes aprenden a trabajar juntos o individualmente para encontrar soluciones apropiadas y únicas a problemas del mundo real que puedan enfrentar, utilizando métodos probados y probados. Más importante aún, se les desafía a pensar tanto de manera creativa como crítica a medida que enfrentan cada problema.

CPS también puede ser influenciada por factores externos como la habilidad individual para lograr metas requeridas a través de un proceso creativo para encontrar nuevas soluciones. La importancia de la comunicación en el proceso educativo significa que los maestros también deben poseer diversas competencias como personalidad, comunicación, sociales, aprendizaje permanente, metodología, planificación, organización, liderazgo y evaluación, para discernir los problemas más significativos para sus encuestados (Suryanto, Degeng, Djatmika & Kuswandi, 2021). La investigación indica que la resolución creativa de problemas puede proporcionar a los estudiantes las habilidades para abordar problemas cotidianos (Abdulla Alabbasi et al., 2021). Estas habilidades requieren práctica extensiva que involucra el proceso creativo, y estas actividades son cruciales para desarrollar habilidades sociales en el campo de la creatividad. Evaluar ideas e involucrar a múltiples personas en la toma de decisiones con pensamiento creativo en la vida cotidiana: el proceso de generar nuevas ideas y aún discutir diferentes formas de pensar.

La integración de CPS y TIC representa intersecciones cruciales en la educación contemporánea. En contextos educativos, las TIC sirven como un conjunto de herramientas poderosas que no solo mejoran los métodos de aprendizaje tradicionales, sino que también fomentan las habilidades de CPS entre los estudiantes (Guillén-Gámez et al., 2024; Mäkiö et al., 2022). Además, las TIC permiten experiencias de aprendizaje personalizadas adaptadas a las necesidades de los estudiantes (Gaeta, Miranda, Orciuoli, Paolozzi & Poce, 2013), capacitándolos para desarrollar habilidades de CPS de manera diversa y atractiva (Andrews-Todd et al., 2023; Treffinger, 2007). A medida que evolucionan los paradigmas educativos, la integración de CPS y TIC no solo prepara a los estudiantes para los desafíos del mundo moderno, sino que también los equipa con habilidades esenciales para prosperar en una sociedad impulsada digitalmente.

Investigaciones recientes subrayan la importancia de integrar las TIC para mejorar las habilidades de CPS entre los estudiantes en entornos educativos. Según Selfa-Sastre, Pifarre, Cujba, Cutillas & Falguera (2022), las TIC desempeñan un papel crucial en la promoción de la creatividad a través de la resolución colaborativa de problemas y la expresión creativa en la educación lingüística. Su estudio destaca cómo las tecnologías digitales permiten diversas oportunidades de aprendizaje y facilitan tres roles clave en el

fomento de la creatividad colaborativa. Estos roles implican el uso de tecnologías interactivas para involucrar a los estudiantes en experiencias de aprendizaje co-creativas, equipándolos con competencias esenciales para enfrentar desafíos complejos en un mundo globalizado e interconectado. Además, Wheeler, Waite, & Bromfield (2002) enfatizan que las TIC permiten a los estudiantes participar en tareas complejas de resolución de problemas que requieren creatividad. Sus hallazgos sugieren que la integración de herramientas de TIC en las aulas educativas no solo mejora las habilidades técnicas de los estudiantes, sino que también cultiva su capacidad para pensar creativamente y abordar problemas desde diferentes ángulos. Estudios como estos destacan colectivamente la sinergia entre CPS y TIC en la educación, destacando las TIC como un catalizador para fomentar el pensamiento creativo y las habilidades de resolución de problemas esenciales para los estudiantes del siglo XXI. Integrar las TIC de manera efectiva en las prácticas pedagógicas no solo enriquece las experiencias educativas, sino que también prepara a los estudiantes para prosperar en un mundo cada vez más complejo y digital.

Varios investigadores han desarrollado instrumentos para evaluar la habilidad de CPS. Por ejemplo, el estudio realizado por Hao et al. (2017) se centró en desarrollar una evaluación estandarizada de habilidades de CPS. Los investigadores reconocieron la importancia de CPS en los entornos de trabajo colaborativos de hoy en día y se propusieron abordar los desafíos prácticos asociados con la evaluación de este constructo complejo. El estudio también destacó la importancia de establecer rúbricas de puntuación claras y criterios para evaluar el desempeño de CPS. En otro estudio, Harding et al. (2017) se centraron en medir CPS utilizando tareas basadas en matemáticas. El estudio destacó el potencial de las tareas basadas en matemáticas para evaluar habilidades de CPS. Emplearon análisis psicométricos rigurosos para examinar la confiabilidad y validez del instrumento de evaluación.

Estos instrumentos desarrollados por diferentes investigadores proporcionan recursos valiosos para evaluar las capacidades de CPS. Ofrecen un enfoque integral para medir diversos aspectos de CPS, incluyendo el pensamiento creativo, estrategias de resolución de problemas y colaboración. Utilizando estos instrumentos, los investigadores y educadores pueden obtener información sobre las habilidades de CPS individuales y adaptar estrategias de instrucción para mejorar las habilidades de resolución creativa de problemas de los estudiantes.

1.2. Medición Rasch

La medición Rasch es un enfoque psicométrico desarrollado por Georg Rasch en la década de 1960 (Panayides et al., 2010). Se utiliza para analizar e interpretar datos de evaluaciones educativas y psicológicas. El modelo Rasch, también conocido como modelo de medición Rasch o modelo Rasch para la teoría de respuesta al ítem (IRT), es un modelo matemático que relaciona la probabilidad de una respuesta a un ítem con el nivel de habilidad o rasgo del individuo evaluado (Cappelleri et al., 2014; Rusch et al., 2017).

El modelo Rasch se basa en el principio de medición probabilística, lo que significa que evalúa la probabilidad de un patrón de respuesta particular dado la habilidad de la persona y la dificultad del ítem (Kyngdon, 2008). Las personas con habilidades más altas deberían tener una mayor probabilidad de responder correctamente a los ítems, reflejando niveles de dificultad más fáciles (Tesio et al., 2023). En otras palabras, las probabilidades están estrechamente relacionadas con las diferencias entre la dificultad del ítem y la habilidad

individual (Boone et al., 2014). El modelo asume que la probabilidad de una respuesta correcta sigue una función logística y que la dificultad del ítem y la habilidad de la persona pueden colocarse en el mismo continuo subyacente, a menudo denominado escala de logit. En la medición Rasch, las habilidades de las personas y las dificultades de los ítems se calibran en una escala intervalar llamada logit, y los parámetros de ítem y persona son completamente independientes (Chan et al., 2021). Esto significa que la medición de las habilidades de los estudiantes permanece igual independientemente del nivel de dificultad de los ítems, y las dificultades de los ítems permanecen invariables independientemente de las habilidades de los estudiantes o los examinados.

La medición Rasch proporciona varias ventajas. Permite el desarrollo de medidas lineales que son independientes de los ítems específicos utilizados en la evaluación (Caly et al., 2008). Esto significa que los puntajes obtenidos de diferentes conjuntos de ítems pueden compararse y agregarse. La medición Rasch también proporciona información sobre la confiabilidad de la medición y el ajuste de los datos al modelo, lo que ayuda a evaluar la calidad del instrumento de evaluación. En la investigación educativa y psicológica, la medición Rasch se utiliza comúnmente para evaluar la calidad de los ítems de prueba, calibrar la dificultad de los ítems, estimar la habilidad de las personas y realizar análisis de ítems y personas. Tiene aplicaciones en diversos campos, incluyendo la evaluación educativa, la investigación de resultados de salud y las ciencias sociales (Planinic et al., 2019). Al emplear el modelo Rasch, los investigadores pueden obtener información valiosa sobre la relación entre individuos e ítems, refinar instrumentos de medición y hacer inferencias significativas sobre el constructo que se está midiendo.

2. Metodología

2.1. Participantes

En este estudio transversal, participaron un total de 137 estudiantes de educación superior como encuestados. Estos estudiantes fueron seleccionados del Departamento de Educación Matemática en Indonesia utilizando una técnica de muestreo aleatorio estratificado. La aprobación ética está siendo considerada por los Comités de Revisión Institucional de la Universitas Islam Negeri Raden Intan Lampung. Esta técnica de muestreo fue elegida para asegurar que la población de la muestra represente con precisión a toda la población investigada. La edad promedio de los participantes fue de 20.84 años, con una desviación estándar (DE) de 1.34. En cuanto a la distribución por género, el 51.8% de los encuestados eran mujeres, mientras que el 48.2% eran hombres. En cuanto a su lugar de residencia, la mayoría de los estudiantes (50.4%) provenían de la ciudad, mientras que el 49.6% restante provenía de otras áreas. Las características de los encuestados se presentan en la Tabla 1.

Tabla 1*Características de los Participantes*

Características	Frecuencia	Porcentaje (%)
Género		
Femenino	71	51.8
Masculino	66	48.2
Grado/Semestre		
4	40	29.2
6	35	25.5
8	62	45.3
Lugar de Residencia		
Ciudad	69	50.4
Urbano	68	49.6

2.2. Instrumento

Los instrumentos utilizados en este estudio fueron desarrollados por investigadores y diseñados específicamente para evaluar las habilidades de resolución creativa de problemas (CPS) de los estudiantes. Estos instrumentos se alinearon con el currículo local de la educación superior para garantizar que midieran de manera efectiva las habilidades y competencias deseadas. Se desarrollaron un total de 20 ítems con este propósito, que abarcan varios aspectos de CPS en el uso de las TIC.

Los ítems en los instrumentos fueron cuidadosamente diseñados para evaluar la capacidad de los estudiantes para aplicar habilidades de pensamiento de orden superior, pensamiento crítico y creativo, estrategias de resolución de problemas y colaboración en el contexto de desafíos del mundo real. Los ítems tenían como objetivo evaluar la capacidad de los estudiantes para generar soluciones innovadoras, pensar críticamente sobre problemas complejos, comunicarse y colaborar eficazmente con otros, y demostrar adaptabilidad y resiliencia en situaciones de resolución de problemas.

Mediante el uso de estos instrumentos, los investigadores buscaban obtener una visión valiosa de las habilidades de CPS de los estudiantes y su capacidad para aplicar estas habilidades en diferentes escenarios. Los instrumentos se desarrollaron para proporcionar una medición confiable y válida de CPS, permitiendo a los investigadores obtener una comprensión integral de las fortalezas y áreas de mejora en este ámbito. El uso de estos instrumentos en este estudio permitió una evaluación sistemática y estandarizada de CPS, proporcionando datos valiosos que pueden contribuir a mejorar las prácticas educativas y el desarrollo curricular.

2.3. Procedimiento

Este estudio implicó un período de recolección de datos de una semana entre estudiantes de educación superior para evaluar sus habilidades de resolución creativa de

problemas (CPS). La prueba de CPS se administró utilizando Google Forms durante las sesiones regulares de clase dedicadas a los cursos respectivos. A los estudiantes se les dio acceso a la prueba a través de su computadora portátil o teléfono móvil y se les concedieron 90 minutos para completarla. En la encuesta de Google Forms, se les pidió a los estudiantes que proporcionaran su información demográfica, incluyendo sexo, lugar de residencia, etnicidad y nivel de grado. La prueba de CPS consistió en ítems de opción múltiple, diseñados para evaluar varios aspectos de las habilidades de CPS. Antes de comenzar la prueba, el investigador proporcionó instrucciones a los estudiantes y presentó tres ejemplos de ítems para familiarizarlos con el formato de las preguntas y los caracteres que aparecerían en la prueba. Una vez respondidas todas las preguntas, los estudiantes enviaron sus respuestas haciendo clic en el botón 'Enviar', lo que guardó sus respuestas para su posterior análisis.

2.4. Análisis de datos

Los datos recolectados en este estudio fueron analizados utilizando la medición Rasch, un enfoque psicométrico ampliamente utilizado para evaluar la concordancia entre los datos observados y el modelo de medición subyacente. En este estudio, se utilizó el programa Winsteps versión 4.7.0 (Linacre, 2020) para analizar los datos, aplicando el modelo Rasch para analizar las respuestas a los ítems del test de CPS. El modelo estima la dificultad de cada ítem y la habilidad de cada estudiante en una escala común de logit.

Los datos obtenidos del instrumento de evaluación de CPS fueron analizados utilizando varios parámetros y técnicas de medición Rasch. Se calcularon el mean square del Outfit (MNSQ) y el Z-standardised del Outfit (ZSTD) para evaluar el ajuste de cada ítem. El Outfit MNSQ proporciona una indicación del grado en que las respuestas observadas se desvían de las respuestas esperadas según el modelo Rasch, mientras que el Outfit ZSTD estandariza los valores de MNSQ para facilitar la comparación entre ítems. Se calculó la correlación punto-medida (Pt-measure correlation) para examinar la relación entre la dificultad del ítem y la habilidad de la persona. Este coeficiente de correlación mide la fuerza de la asociación entre las respuestas a los ítems y las habilidades estimadas de las personas en la escala de logit.

El mapa de Wright, una representación gráfica, se utilizó para mostrar la distribución de las dificultades de los ítems y las habilidades correspondientes de los estudiantes. Este mapa proporciona una visión general de la jerarquía de dificultades de los ítems y el rango de habilidades exhibidas por los estudiantes. Además, se realizó un análisis de valor logit de la persona (LVP) para identificar las habilidades de CPS de los estudiantes.

Para explorar posibles diferencias en el funcionamiento de los ítems según el género y el lugar de residencia, se llevó a cabo un análisis de DIF. El análisis de DIF identifica ítems que pueden funcionar de manera diferente para diferentes grupos, lo que indica un posible sesgo o desempeño diferencial entre grupos.

Para analizar las diferencias en las habilidades de CPS entre los estudiantes, se utilizó el software SPSS versión 26. Se calcularon estadísticas descriptivas como la media y la desviación estándar (SD) para proporcionar una visión general de los datos. Además, se emplearon estadísticas del paquete R para visualizar el mapa.

3. Resultados

3.1. Pregunta de investigación 1: ¿El instrumento desarrollado demuestra fiabilidad y validez basadas en la medición de Rasch?

Los resultados del análisis de validación utilizando la medición Rasch se presentan en la Tabla 2.

Tabla 2

Resultados del Análisis Rasch realizado en CPS

Características	Ítem	Persona
Número de Ítems	20	137
infit MNSQ		
Media	1.00	1.00
SD	0.18	0.14
outfit MNSQ		
Media	1.02	1.02
SD	0.30	0.29
Separación	4.01	1.40
Confiabilidad	0.66	0.94
Varianza bruta explicada por las medidas	76.6%	

3.2. Pregunta de investigación 2: ¿Cuáles son los patrones de interacción entre ítems y personas en el instrumento desarrollado según el mapa de Wright?

El patrón de interacción entre los ítems y los individuos en el instrumento desarrollado, basado en el mapa de Wright, se presenta en la Figura 1. En la Figura 1 se observa que el instrumento consta de 20 ítems e involucra a 137 estudiantes como encuestados. La línea vertical del lado derecho representa los ítems, mientras que el lado izquierdo representa el número de encuestados. Se puede notar que el ítem número 13 (Q13) cae en la categoría de ítems fáciles, mientras que el ítem número 12 (Q12) se clasifica como un ítem difícil. La distribución o características de los ítems difíciles y fáciles se pueden observar en la Figura 2. Por otro lado, la distribución del ajuste de los ítems se muestra en la Figura 3.

Figura 1

Mapa de Wright

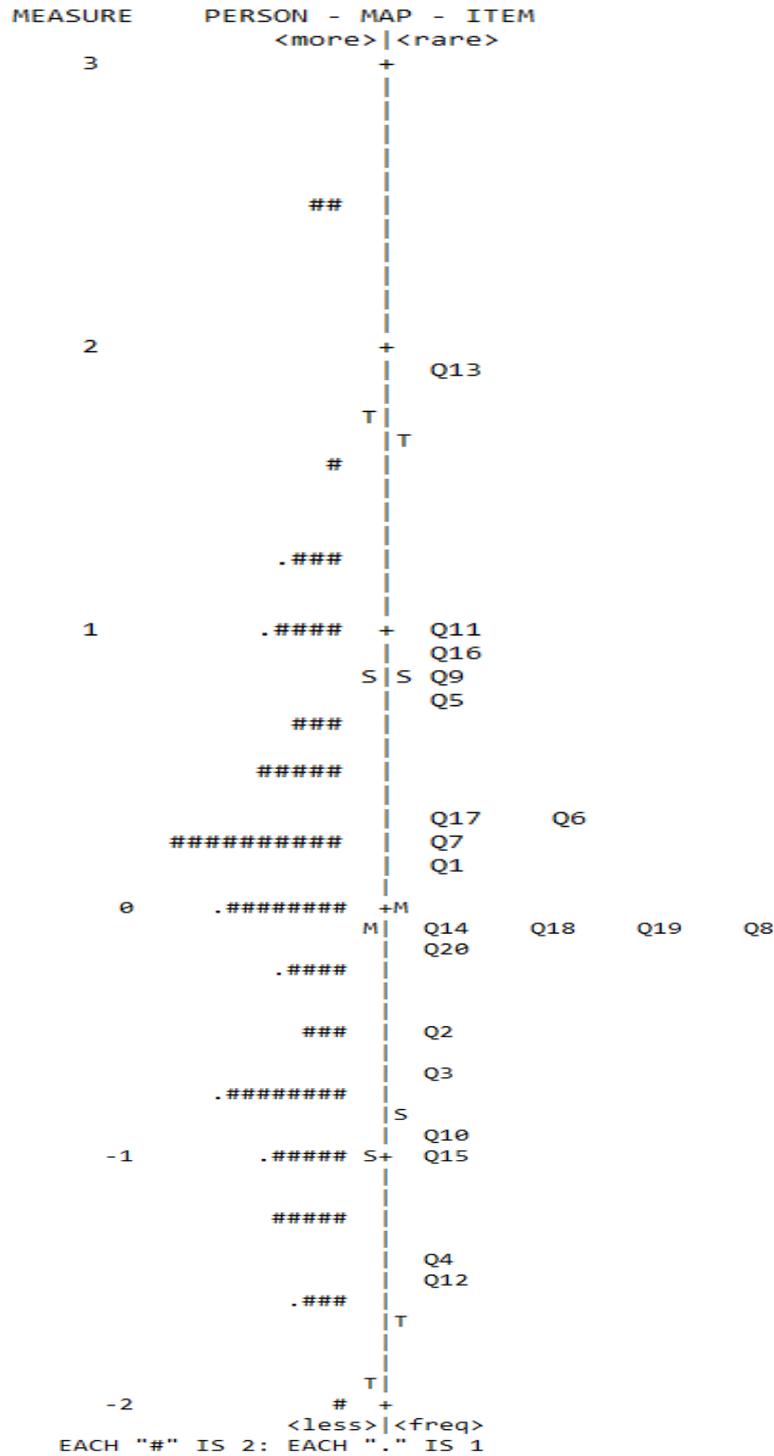


Figura 2

Un ítem pertenece a los ítems difíciles y fáciles

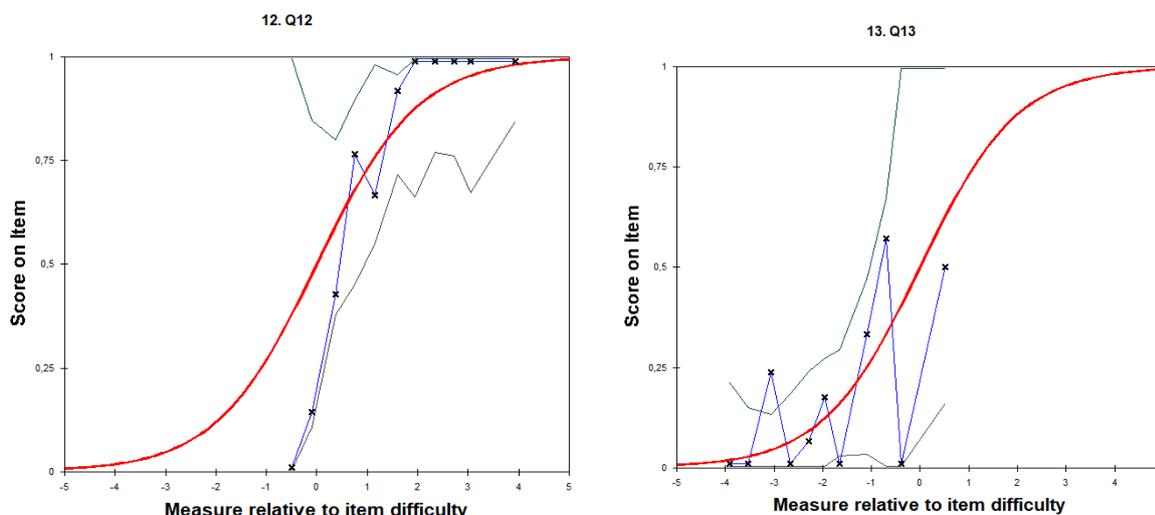
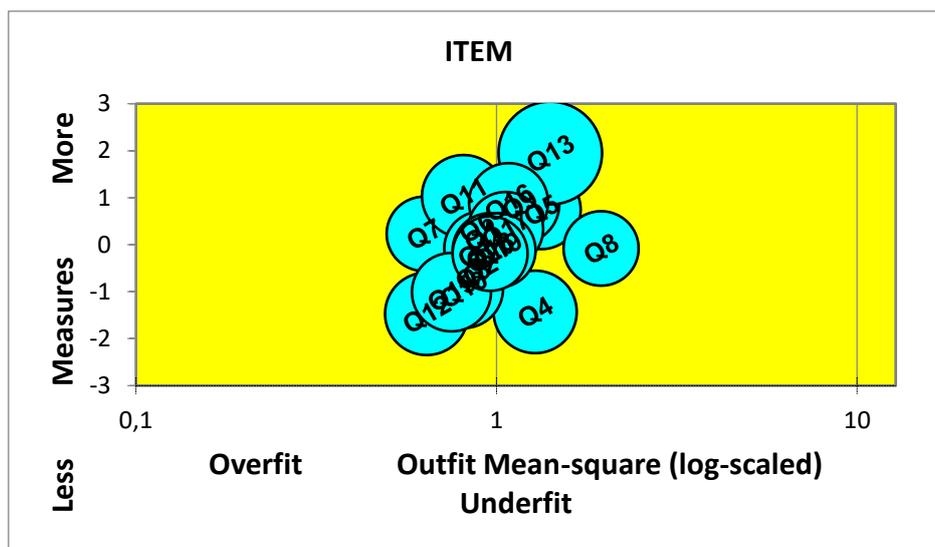


Figura 3

La distribución de los ítems se basa en el Mapa de Burbujas



Para determinar el ajuste de los ítems desarrollados basados en el modelo Rasch, se consideraron tres criterios: Outfit MNSQ, Outfit ZSTD y Pt-Measure Corr. Un rango entre 0.5 y 1.5 para los valores de Outfit MNSQ tanto para ítems como para individuos indica un buen ajuste entre los datos y el modelo. Los valores de Outfit ZSTD entre -1.9 y 1.9 implican que los ítems pueden predecirse. Además, Pt-Measure Corr se utiliza para determinar si los ítems miden el constructo deseado. Si el valor es positivo (+), indica que el ítem mide el constructo deseado. Por el contrario, si el valor es negativo (-), el ítem no mide el constructo deseado.

Tabla 3*Distribución de Ítems Basada en 3 Criterios de Ajuste de Ítem*

Ítems	Medida	Infit MNSQ	Outfit MNSQ	Outfit SZTD	Correlación Pt-Measure
8	-0.08	1.60	1.95	8.51	-0.27
13	1.95	1.05	1.41	1.47	0.22
5	0.74	1.11	1.33	2.52	0.22
4	-1.43	1.15	1.28	1.47	0.16
9	0.86	1.09	1.16	1.17	0.27
16	0.90	1.16	1.08	0.63	0.24
17	0.33	1.04	1.06	0.67	0.34
19	-0.12	1.05	1.01	0.19	0.35
20	-0.19	1.02	0.96	-0.41	0.38
18	-0.12	1.00	0.96	-0.43	0.40
6	0.37	0.95	0.88	-1.26	0.45
14	-0.08	0.94	0.91	-1.05	0.45
1	0.19	0.91	0.91	-1.00	0.48
3	-0.64	0.91	0.86	-1.30	0.47
2	-0.53	0.89	0.90	-0.98	0.47
11	1.03	0.89	0.81	-1.35	0.49
10	-0.94	0.86	0.81	-1.48	0.50
15	-1.01	0.82	0.75	-1.90	0.54
12	-1.47	0.80	0.64	-2.13	0.54
7	0.23	0.69	0.63	-4.57	0.70

Basado en los tres criterios mencionados, es evidente que el ítem 8 (Q8) no cumple con los criterios mencionados anteriormente, lo que indica que el ítem no se ajusta bien. Por lo tanto, se recomienda eliminar o revisar el ítem 8. Como se muestra en la Figura 5, Q8 parece acercarse a un ajuste insuficiente, lo que indica que no se alinea bien con el modelo Rasch.

3.3. Pregunta de investigación 3: ¿Existen sesgos en el instrumento basados en el género según el análisis de Funcionamiento Diferencial del Ítem (DIF)?

Se realizó un análisis de DIF para determinar si había ítems que favorecían a un género (en el contexto de este estudio). Un ítem se considera que tiene DIF si el valor t es menor que -2.0 o mayor que 2.0, el valor de contraste de DIF es menor que -0.5 o mayor que 0.5, y el valor de p es menor que 0.05 o mayor que -0.05 (Bond & Fox, 2015; Boone et al., 2014). Aquí están los resultados del análisis utilizando el modelo Rasch.

Tabla 4

DIF Potencial Debido al Género

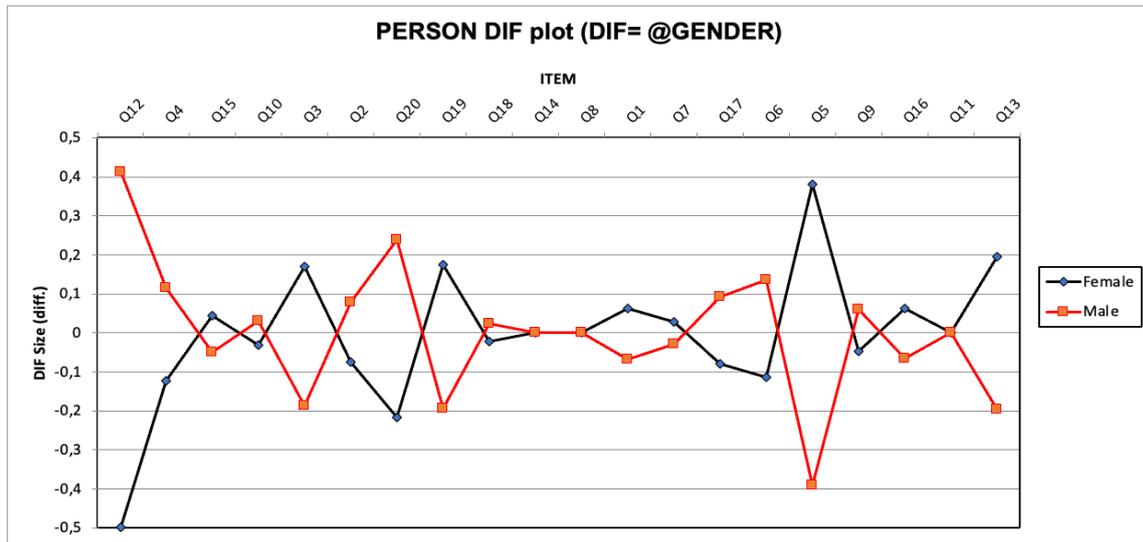
Ítem	DIF		Contraste DIF	t-value	Prob.
	Feminino	Masculino			
Q12	-1.97	-1.06	-0.91	-2.04	0.0432

El análisis revela que el ítem Q12 es un ítem difícil, lo que indica que puede diferenciar las habilidades entre hombres y mujeres. Esto es apoyado por el análisis de DIF, que examina el rendimiento del ítem en los grupos de género. El gráfico de DIF (Fig. 4) proporciona una representación visual de los valores de DIF para cada ítem.

En el gráfico, los valores de DIF para el ítem Q12 son notablemente más altos en comparación con los otros ítems. Esto sugiere que hay una diferencia significativa en el rendimiento de hombres y mujeres en este ítem en particular. El análisis de DIF indica que el ítem Q12 puede favorecer a un género sobre el otro en términos de dificultad o discriminación. Estos hallazgos son importantes porque destacan posibles sesgos relacionados con el género en la medición de habilidades colaborativas para resolver problemas. Puede ser necesario realizar una investigación adicional y una revisión potencial del ítem para garantizar una evaluación justa e imparcial de todas las personas, independientemente de su género.

Figura 4

Potencial DIF Debido al Género



3.4. Pregunta de investigación 4: ¿Cómo se desarrolla la resolución colaborativa de problemas (CPS) en los estudiantes en términos de calificaciones del curso?

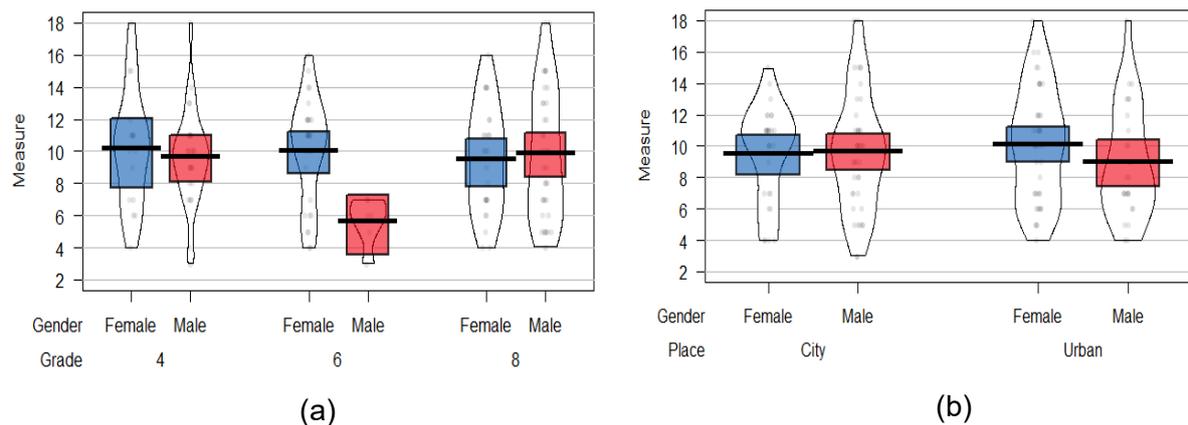
La descripción estadística de las respuestas de los estudiantes a los ítems dados se presenta en la Tabla 9. En la Figura 5 (a), se puede observar que entre las estudiantes mujeres, aquellas en el cuarto semestre tienen una habilidad de respuesta promedio (M) de $M = 10.19$, $SD = 3.73$, seguidas por las del sexto semestre con $M = 10.00$, $SD = 3.32$, y las

del octavo semestre con $M = 9.50$, $SD = 3.37$. Por otro lado, entre los estudiantes hombres, aquellos en el 4º semestre tienen una habilidad de respuesta promedio de $M = 9.67$, $SD = 3.11$, seguidos por los del 6º semestre con $M = 5.67$, $SD = 1.51$, y los del 8º semestre con $M = 9.86$, $SD = 4.02$.

Sin embargo, al comparar las habilidades de los estudiantes según su lugar de residencia (urbano versus rural), existen diferencias como se muestra en la Figura 5 (b). Las estudiantes mujeres que residen en áreas urbanas tienen una habilidad de respuesta promedio de $M = 9.52$, $SD = 3.11$, mientras que las de áreas rurales tienen $M = 10.10$, $SD = 3.59$. De manera similar, los estudiantes hombres que residen en áreas urbanas tienen una habilidad de respuesta promedio de $M = 9.68$, $SD = 3.81$, mientras que los de áreas rurales tienen $M = 9.00$, $SD = 3.58$. Estos hallazgos indican que la ubicación de residencia puede influir en cierta medida en las habilidades colaborativas para resolver problemas de los estudiantes.

Figura 5

La capacidad de los estudiantes para responder según el grado y el género (a), y el grado y el lugar de residencia (b)



4. Discusión

En general, este análisis proporciona una comprensión de las características de medición de los ítems e individuos en este estudio. Los hallazgos indican que la medición utilizada se ajusta razonablemente bien al modelo de medición Rasch, con la capacidad de diferenciar niveles de dificultad entre los ítems y un nivel de confiabilidad suficientemente alto. El respaldo de investigaciones relevantes en este campo también confirma estos hallazgos y proporciona una base sólida para entender las características de medición de este estudio. Estudios previos, como el realizado por Soeharto (2021), también encontraron resultados similares en términos de ajuste, separación y confiabilidad de la medición.

Los resultados del análisis que indican un buen ajuste entre los ítems e individuos con el modelo de medición Rasch sirven como indicador de que la medición representa con precisión las características medidas. La exitosa separación de niveles de dificultad entre los ítems también proporciona una ventaja al proporcionar información más detallada y precisa sobre las habilidades individuales medidas. Esto es consistente con investigaciones

previas que afirman que una separación adecuada es crucial para garantizar una medición confiable y válida (Soeharto & Csapó, 2022).

Además, la confiabilidad razonablemente buena tanto para ítems como para individuos proporciona la seguridad de que los resultados de la medición obtenidos son confiables y consistentes. En el contexto de este estudio, una confiabilidad de 0.66 para ítems y 0.94 para personas indica un nivel satisfactorio de confiabilidad. Otros estudios relevantes también pueden proporcionar apoyo para el análisis de las características de medición realizadas en este estudio. Por ejemplo, un estudio realizado por Chan et al. (2021) encontró resultados similares en términos de ajuste y confiabilidad de la medición. Además, la investigación de Avinç & Doğan (2024) encontró que la validez y confiabilidad fueron confirmadas por el modelo Rasch. Sin embargo, señalaron que sería ventajoso probar su validez y confiabilidad en diversas clases, grupos de edad y niveles educativos. Asimismo, este estudio, al igual que el de Welter et al. (2024) muestra que las propiedades psicométricas son válidas y confiables utilizando la medición Rasch. En general, este análisis proporciona una comprensión profunda de las características de medición de los ítems e individuos en este estudio. El respaldo de investigaciones relevantes y los resultados del análisis que muestran buen ajuste, separación y confiabilidad brindan confianza en que la medición realizada en este estudio es confiable y proporciona información válida sobre las características medidas.

Basándose en el patrón de interacción observado en la Figura 3, se pueden sacar conclusiones sobre el nivel de dificultad de los ítems. Por ejemplo, se observa que el ítem 13 (Q13) está posicionado más bajo en la línea vertical, indicando que pertenece a la categoría de ítems fáciles. Por el contrario, el ítem 12 (Q12) está posicionado más alto, indicando que pertenece a la categoría de ítems difíciles.

Este análisis proporciona información importante sobre el nivel de dificultad de cada elemento en el instrumento desarrollado. Cuando se conoce el nivel de dificultad de los ítems, pueden hacerse ajustes y desarrollos adicionales para asegurar que los ítems utilizados cubran niveles de dificultad apropiados alineados con los objetivos de investigación. Sin embargo, es importante destacar que la evaluación del nivel de dificultad de los ítems no se basa únicamente en la posición de los ítems en la línea vertical de la Figura 3, sino que también tiene en cuenta otros factores como las características de los encuestados y el contexto más profundo de la medición. Los ítems fueron efectivos para evaluar las habilidades CPS de los estudiantes utilizando TIC en el aula (Wheeler et al., 2002) y su impacto en las competencias digitales de los estudiantes (Guillén-Gámez et al., 2024) lo que sugiere que el nivel de dificultad de los ítems fue adecuado para el propósito previsto del instrumento (Hobani & Alharbi, 2024). La información sobre la dificultad de los ítems puede guiar la refinación y el desarrollo del instrumento, permitiendo a los investigadores identificar áreas donde el nivel de dificultad puede necesitar ajustes, ya sea modificando los ítems existentes o agregando nuevos ítems para cubrir diferentes niveles de dificultad.

Pasando al análisis de Funcionamiento Diferencial del Ítem (DIF), el DIF se refiere a las diferencias en las características de respuesta a un ítem entre dos o más grupos de encuestados que deberían tener el mismo nivel de habilidad. En este contexto, se exploran las diferencias de habilidad entre hombres y mujeres utilizando el concepto de DIF. En este estudio, se encontró que Q12 tiene el potencial para diferenciar la habilidad entre hombres y mujeres, con una categoría de DIF. Esto indica que hombres y mujeres tienen

probabilidades diferentes de responder Q12 a pesar de tener el mismo nivel de habilidad. En este contexto, hay indicios de que Q12 puede ser más difícil para un grupo de género.

Sin embargo, es importante destacar que el análisis de DIF solo proporciona indicaciones preliminares de posibles diferencias en las características de respuesta entre grupos de encuestados. Es crucial ver estos hallazgos de DIF como información que puede ayudar en el desarrollo del instrumento y obtener una mejor comprensión de cómo los ítems de comportamiento en el instrumento se desempeñan en grupos específicos. En otras palabras, "DIF no es sinónimo de sesgo", como señaló Zieky (2012). Los ítems identificados como DIF no implican necesariamente sesgo. Según Mollazehi & Abdel-Salam (2024), el sesgo se refiere al rendimiento diferencial entre individuos de igual habilidad de diferentes subgrupos debido a factores irrelevantes. El DIF, introducido para distinguir el significado estadístico del sesgo de sus implicaciones sociales, se centra en las propiedades estadísticas diferentes que los ítems exhiben entre los subgrupos después de igualar las habilidades individuales (Angoff, 2012). Dado que la interpretación del DIF se limita a las diferencias en propiedades estadísticas, como la dificultad y la discriminación del ítem, son necesarias revisiones por un panel de expertos para determinar si los ítems con DIF son sesgados (H. Lee & Geisinger, 2014). Por lo tanto, los ítems que muestran DIF pueden incluirse en una prueba si no se encuentra evidencia adecuada de sesgo a través de la revisión del panel (De Ayala et al., 2002). En futuras investigaciones, se pueden tomar medidas para evaluar las causas del DIF y asegurar que el instrumento mida con precisión sin sesgo de género.

El análisis estadístico presentado en la Figura 5 proporciona información sobre la capacidad de los estudiantes para responder a los ítems dados según las categorías de género, semestre y ubicación residencial. Los datos ofrecen una visión general de la variación en las habilidades de respuesta entre diferentes grupos según el género, semestre y ubicación residencial. Las diferencias en medias y desviaciones estándar indican variaciones en la comprensión del material o enfoques de aprendizaje entre estos grupos.

El rendimiento de las estudiantes en el cuarto semestre, que exhiben un puntaje de habilidad promedio más alto en la resolución creativa de problemas mediante el uso de TIC, parece superar al de los estudiantes masculinos en diferentes semestres. Esta observación se alinea con investigaciones previas que indican que las estudiantes a menudo demuestran niveles más altos de competencia en tareas de resolución de problemas que requieren habilidades colaborativas y relacionadas con las TIC (Andrews-Todd et al., 2023; S. W.-Y. Lee et al., 2023). Los estudios han mostrado consistentemente que las mujeres tienden a destacarse en entornos de aprendizaje colaborativo, utilizando eficazmente las herramientas TIC para mejorar sus habilidades de resolución de problemas (Ma et al., 2023). Esta tendencia se atribuye a varios factores, incluyendo una mayor atención al detalle, habilidades mejoradas de comunicación y una preferencia por el trabajo en equipo, que son críticos en tareas de resolución creativa de problemas (Thornhill-Miller et al., 2023).

Sin embargo, al comparar a los estudiantes que viven en áreas urbanas y rurales, se pueden observar diferencias en los resultados. Las estudiantes que viven en áreas rurales muestran el mejor rendimiento en la resolución creativa de problemas mediante el uso de TIC. Las áreas rurales a menudo enfrentan desafíos únicos como el acceso limitado a recursos, incluida la infraestructura educativa y la tecnología (Alabdali et al., 2023). A pesar de estos desafíos, las estudiantes en áreas rurales pueden exhibir habilidades más altas para resolver problemas debido a su adaptabilidad y resiliencia para navegar estas limitaciones. La investigación sugiere que las mujeres a menudo muestran mayores niveles

de persistencia y adaptabilidad en entornos de aprendizaje (Dabas et al., 2023), lo que podría contribuir a su mejor desempeño en tareas de resolución creativa de problemas que involucran TIC. Además, los factores culturales y sociales pueden influir en los resultados educativos (Min, 2023). En algunas culturas, puede haber un énfasis más fuerte en la educación para las mujeres, especialmente en entornos rurales donde el acceso a oportunidades educativas puede considerarse transformador para individuos y familias (Robinson-Pant, 2023). Este énfasis podría motivar a las estudiantes a destacarse académicamente y en tareas de resolución de problemas, incluyendo aquellas que involucran TIC.

Estos hallazgos destacan la influencia potencial del género, semestre y ubicación residencial en las habilidades de respuesta de los estudiantes. Las variaciones en las medias y desviaciones estándar sugieren diferencias en las experiencias de aprendizaje, la exposición a recursos educativos u otros factores contextuales que pueden contribuir a las variaciones en las habilidades de respuesta.

En general, estos análisis proporcionan información valiosa sobre las características de la medición, el funcionamiento de los elementos diferenciales y la relación entre las habilidades de respuesta y diversos factores como el género, semestre y ubicación residencial. Contribuyen a una mejor comprensión de los datos y ofrecen implicaciones para futuras investigaciones y el desarrollo de instrumentos en este campo.

5. Limitaciones y Futuras Investigaciones

Este estudio proporciona contribuciones importantes al desarrollo de instrumentos de evaluación para evaluar a estudiantes en CPS. Sin embargo, existen algunas limitaciones que deben abordarse. En primer lugar, la fiabilidad de la medición para los ítems obtuvo un valor de 0.66, lo que indica un nivel moderado de fiabilidad. Aunque esta fiabilidad puede ser aceptable en algunos contextos de investigación, es deseable mejorarla para el desarrollo de instrumentos de evaluación más robustos en el futuro.

Así mismo, existe un ítem, el ítem 8 (Q8), que no cumple con los criterios en el análisis de ajuste del ítem con el modelo Rasch. Este ítem podría ser eliminado o revisado para asegurar un mejor ajuste y validez del instrumento de evaluación. La revisión y refinamiento de los ítems que no cumplen con los criterios es necesaria para asegurar que el instrumento de evaluación produzca resultados más precisos y confiables.

Además, se observa un potencial funcionamiento diferencial del ítem (DIF) basado en género en el ítem Q12. Esto indica un sesgo del instrumento hacia el género en términos de responder preguntas difíciles. En el desarrollo de futuros instrumentos de evaluación, es importante abordar este sesgo para asegurar un instrumento más neutral y justo para todos los participantes.

El estudio también proporciona información sobre los patrones de interacción entre los ítems y los individuos en el instrumento desarrollado basado en el mapa de Wright. Al examinar estos patrones, se puede entender el nivel de dificultad de cada ítem y la distribución de las habilidades de los encuestados. Sin embargo, no se proporciona una explicación adicional de las implicaciones de estos patrones de interacción en el desarrollo de instrumentos de evaluación.

A pesar de estas limitaciones, este estudio sienta una sólida base para futuras investigaciones en el desarrollo de instrumentos de evaluación basados en habilidades de pensamiento de orden superior (HOTS). Las investigaciones futuras pueden centrarse en mejorar la fiabilidad del instrumento, eliminar el sesgo del instrumento basado en género y explorar más a fondo los patrones de interacción entre los ítems y los individuos.

En estudios futuros, es importante involucrar una muestra más representativa y ampliar el alcance del análisis para obtener resultados más generalizables. Además, la validación del instrumento también puede realizarse utilizando otros métodos que proporcionen información adicional sobre el ajuste del instrumento, su validez y fiabilidad.

En resumen, este estudio contribuye al desarrollo de instrumentos de evaluación basados en HOTS utilizando el modelo Rasch. A pesar de algunas limitaciones que deben abordarse, este estudio sirve como una base importante para futuras investigaciones en el desarrollo de instrumentos de evaluación más efectivos y robustos.

6. Conclusiones

Basándose en el análisis realizado, se pueden extraer las siguientes conclusiones: La medición utilizando el modelo Rasch muestra un buen ajuste entre los ítems y los individuos en el instrumento de evaluación. Los ítems muestran un buen ajuste con el modelo Rasch, permitiendo diferenciar niveles de dificultad entre diferentes ítems, y también tienen un nivel de fiabilidad razonablemente bueno. Existe un patrón de interacción entre los ítems y los individuos en el instrumento de evaluación basado en el mapa de Wright. Los ítems en el instrumento pueden diferenciar efectivamente las habilidades de los individuos, siendo algunos ítems relativamente fáciles y otros más desafiantes. Hay ítems que no cumplen con los criterios de Funcionamiento Diferencial del Ítem (DIF) basados en género. El ítem Q12 en el instrumento de evaluación tiende a favorecer a los hombres sobre las mujeres en términos de habilidad para responder. Las estudiantes mujeres en el cuarto semestre tienen habilidades promedio de respuesta más altas en comparación con las estudiantes mujeres en el sexto y octavo semestre. Sin embargo, existen diferencias en las habilidades de respuesta entre estudiantes hombres y mujeres, así como entre estudiantes que viven en áreas urbanas y rurales.

Estas conclusiones indican que el desarrollo de instrumentos de evaluación para CPS puede arrojar resultados de medición válidos y confiables. No obstante, se debe tener en cuenta que hay algunos ítems que deben mejorarse para lograr una mayor precisión. Además, existe una indicación de sesgo del instrumento basado en género en términos de habilidades de respuesta. Esto debe considerarse al desarrollar instrumentos que sean más neutrales en cuanto al género y justos en la medición de las habilidades de los participantes.

En el desarrollo de instrumentos de evaluación para CPS y el uso del modelo Rasch, hay impactos positivos en el desarrollo del currículo y la instrucción. Es importante que los educadores adopten enfoques efectivos de diferenciación, aseguren instrumentos de evaluación neutrales en cuanto al género y consideren factores contextuales en el diseño de una instrucción inclusiva y centrada en el aprendiz que se alinee con el potencial de los estudiantes. De este modo, la educación puede volverse más relevante, receptiva y capacitar a los estudiantes para enfrentar los desafíos de un mundo complejo.

Authors' Contribution

Farida Farida: Conceptualization, Writing - Original Draft, Formal analysis; Yosep Aspat Alamsyah: Methodology, Editing, and Visualization; Bambang Sri Anggoro: Supervision, Funding acquisition, Writing – review & editing; Tri Andari: Formal analysis and Visualization; Restu Lusiana: Editing and Formal analysis.

Financiación

El Departamento de Investigación y Servicio Comunitario (LP2M) financió el estudio reportado en la Universitas Islam Negeri Raden Intan Lampung, Indonesia (Número de concesión 99/2023).

Referencias

- Abdulla Alabbasi, A. M., Hafsyah, A. S., Runco, M. A., & AlSaleh, A. (2021). Problem finding, divergent thinking, and evaluative thinking among gifted and nongifted students. *Journal for the Education of the Gifted*, 44(4), 398–413. <https://doi.org/10.1515/ctra-2018-0019>
- Abina, A., Temeljotov Salaj, A., Cestnik, B., Karalič, A., Ogrinc, M., Kovačič Lukman, R., & Zidanšek, A. (2024). Challenging 21st-Century Competencies for STEM Students: Companies' Vision in Slovenia and Norway in the Light of Global Initiatives for Competencies Development. *Sustainability*, 16(3), 1295. <https://doi.org/10.3390/su16031295>
- Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education*, 4(1), 1–11. <https://doi.org/10.11648/j.ijsedu.20160401.11>
- Alabdali, S. A., Pileggi, S. F., & Cetindamar, D. (2023). Influential factors, enablers, and barriers to adopting smart technology in rural regions: A literature review. *Sustainability*, 15(10), 7908. <https://doi.org/10.3390/su15107908>
- Andrews-Todd, J., Jiang, Y., Steinberg, J., Pugh, S. L., & D'Mello, S. K. (2023). Investigating collaborative problem solving skills and outcomes across computer-based tasks. *Computers & Education*, 207, 104928. <https://doi.org/10.1016/j.compedu.2023.104928>
- Angoff, W. H. (2012). Perspectives on differential item functioning methodology. In *Differential item functioning* (pp. 3–23). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9780203357811-2/perspectives-differential-item-functioning-methodology-william-angoff>
- Arredondo-Trapero, F. G., Guerra-Leal, E. M., Kim, J., & Vázquez-Parra, J. C. (2024). Competitiveness, quality education and universities: The shift to the post-pandemic world. *Journal of Applied Research in Higher Education*. <https://doi.org/10.1108/JARHE-08-2023-0376>
- Avinç, E., & Doğan, F. (2024). Digital literacy scale: Validity and reliability study with the rasch model. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12662-7>
- Bond, T., & Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences, Third Edition* (3rd ed.). Routledge. <https://doi.org/10.4324/9781315814698>
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer. <https://doi.org/10.1007/978-94-007-6857-4>

- Brunner, L. G., Peer, R. A. M., Zorn, C., Paulik, R., & Logan, T. M. (2024). Understanding cascading risks through real-world interdependent urban infrastructure. *Reliability Engineering & System Safety*, 241, 109653. <https://doi.org/10.1016/j.ress.2023.109653>
- Burns, D. P., & Norris, S. P. (2009). Open-minded environmental education in the science classroom. *Paideusis*, 18(1), 36–43. <https://doi.org/10.7202/1072337ar>
- Caena, F., & Redecker, C. (2019). Aligning teacher competence frameworks to 21st century challenges: The case for the European Digital Competence Framework for Educators (Digcompedu). *European Journal of Education*, 54(3), 356–369. <https://doi.org/10.1111/ejed.12345>
- Cappelleri, J. C., Lundy, J. J., & Hays, R. D. (2014). Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. *Clinical Therapeutics*, 36(5), 648–662. <https://doi.org/10.1016/j.clinthera.2014.04.006>
- Care, E., & Kim, H. (2018). Assessment of twenty-first century skills: The issue of authenticity. *Assessment and Teaching of 21st Century Skills: Research and Applications*, 21–39. https://doi.org/10.1007/978-3-319-65368-6_2
- Carnevale, A. P., & Smith, N. (2013). Workplace basics: The skills employees need and employers want. In *Human Resource Development International* (Vol. 16, Issue 5, pp. 491–501). Taylor & Francis. <https://doi.org/10.1080/13678868.2013.821267>
- Caty, G. D., Arnould, C., Stoquart, G. G., Thonnard, J.-L., & Lejeune, T. M. (2008). ABILOCO: a Rasch-built 13-item questionnaire to assess locomotion ability in stroke patients. *Archives of Physical Medicine and Rehabilitation*, 89(2), 284–290. <https://doi.org/10.1016/j.apmr.2007.08.155>
- Chan, S. W., Looi, C. K., & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: A rasch model measurement analysis. *Journal of Computers in Education*, 8(2), 213–236. <https://doi.org/10.1007/s40692-020-00177-2>
- Cho, S., & Lin, C.-Y. (2010). Influence of family processes, motivation, and beliefs about intelligence on creative problem solving of scientifically talented individuals. *Roeper Review*, 33(1), 46–58. <https://doi.org/10.1080/02783193.2011.530206>
- Dabas, C. S., Muljana, P. S., & Luo, T. (2023). Female students in quantitative courses: An exploration of their motivational sources, learning strategies, learning behaviors, and course achievement. *Technology, Knowledge and Learning*, 28(3), 1033–1061. <https://doi.org/10.1007/s10758-021-09552-z>
- De Ayala, R. J., Kim, S.-H., Stapleton, L. M., & Dayton, C. M. (2002). Differential Item Functioning: A Mixture Distribution Conceptualization. *International Journal of Testing*, 2(3–4), 243–276. <https://doi.org/10.1080/15305058.2002.9669495>
- Farida, F., Supriadi, N., Andriani, S., Pratiwi, D. D., Suherman, S., & Muhammad, R. R. (2022). STEM approach and computer science impact the metaphorical thinking of Indonesian students'. *Revista de Educación a Distancia (RED)*, 22(69). <https://doi.org/10.6018/red.493721>
- Florida, R. (2014). The creative class and economic development. *Economic Development Quarterly*, 28(3), 196–205. <https://doi.org/10.1177/0891242414541693>

- Gaeta, M., Miranda, S., Orciuoli, F., Paolozzi, S., & Poce, A. (2013). An Approach To Personalized e-Learning. *Journal of Education, Informatics & Cybernetics*, 11(1). <https://www.iisc.org/Journal/pdv/sci/pdfs/HEB785ZU.pdf>
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013). Complex problem solving in educational contexts—Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, 105(2), 364. <https://doi.org/10.1037/a0031856>
- Guillén-Gámez, F. D., Gómez-García, M., & Ruiz-Palmero, J. (2024). Digital competence in research work: Predictors that have an impact on it according to the type of university and gender of the Higher Education teacher:[Digital competence in research work: predictors that have an impact on it according to the type of university and gender of the Higher Education teacher]. *Pixel-Bit. Revista de Medios y Educación*, 69, 7–34. <https://doi.org/10.12795/pixelbit.99992>
- Hao, J., Liu, L., von Davier, A. A., & Kyllonen, P. C. (2017). Initial steps towards a standardized assessment for collaborative problem solving (CPS): Practical challenges and strategies. *Innovative Assessment of Collaboration*, 135–156. https://doi.org/10.1007/978-3-319-33261-1_9
- Harding, S.-M. E., Griffin, P. E., Awwal, N., Alom, B. M., & Scoular, C. (2017). Measuring collaborative problem solving using mathematics-based tasks. *AERA Open*, 3(3), 2332858417728046. <https://doi.org/10.1177/2332858417728046>
- Hobani, F., & Alharbi, M. (2024). A Psychometric Study of the Arabic Version of the “Searching for Hardships and Obstacles to Shots (SHOT)” Instrument for Use in Saudi Arabia. *Vaccines*, 12(4), 391. <https://doi.org/10.3390/vaccines12040391>
- Küçükaydın, M. A., Çite, H., & Ulum, H. (2024). Modelling the relationships between STEM learning attitude, computational thinking, and 21st century skills in primary school. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12492-7>
- Kyngdon, A. (2008). The Rasch model from the perspective of the representational theory of measurement. *Theory & Psychology*, 18(1), 89–109. <https://doi.org/10.1177/095935430708692>
- Lee, D., & Lee, Y. (2024). Productive failure-based programming course to develop computational thinking and creative Problem-Solving skills in a Korean elementary school. *Informatics in Education*, 23(2), 385–408. <https://doi.org/10.1007/s10956-024-10130-y>
- Lee, H., & Geisinger, K. F. (2014). The Effect of Propensity Scores on DIF Analysis: Inference on the Potential Cause of DIF. *International Journal of Testing*, 14(4), 313–338. <https://doi.org/10.1080/15305058.2014.922567>
- Lee, S. W.-Y., Liang, J.-C., Hsu, C.-Y., Chien, F. P., & Tsai, M.-J. (2023). Exploring Potential Factors to Students’ Computational Thinking. *Educational Technology & Society*, 26(3), 176–189. <https://doi.org/10.30191/ETS>
- Lee, T., O’Mahony, L., & Lebeck, P. (2023). Creative Problem-Solving. In T. Lee, L. O’Mahony, & P. Lebeck, *Creativity and Innovation* (pp. 117–147). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-8880-6_5

- Lestari, W., Sari, M. M., Istiyadi, M., & Fahmi, F. (2023). Analysis of Implementation of the Independent Curriculum in Science Learning at SMP Negeri 1 Tanah Grogot Kalimantan Timur, Indonesia. *Repository Universitas Lambung Mangkurat*. <https://doi.org/10.36348/jaep.2023.v07i06.001>
- Linacre, J. M. (2020). *Winsteps®(Version 4.7. 0)[Computer Software].(4.7. 0). Winsteps. Com*.
- Liu, X., Gu, J., & Xu, J. (2024). The impact of the design thinking model on pre-service teachers' creativity self-efficacy, inventive problem-solving skills, and technology-related motivation. *International Journal of Technology and Design Education*, 34(1), 167–190. <https://doi.org/10.1007/s10798-023-09809-x>
- Lorusso, L., Lee, J. H., & Worden, E. A. (2021). Design thinking for healthcare: Transliterating the creative problem-solving method into architectural practice. *HERD: Health Environments Research & Design Journal*, 14(2), 16–29. <https://doi.org/10.1177/193758672199422>
- Ma, Y., Zhang, H., Ni, L., & Zhou, D. (2023). Identifying collaborative problem-solver profiles based on collaborative processing time, actions and skills on a computer-based task. *International Journal of Computer-Supported Collaborative Learning*, 18(4), 465–488. <https://doi.org/10.1007/s11412-023-09400-5>
- Mäkiö, E., Azmat, F., Ahmad, B., Harrison, R., & Colombo, A. W. (2022). T-CHAT educational framework for teaching cyber-physical system engineering. *European Journal of Engineering Education*, 47(4), 606–635. <https://doi.org/10.1080/03043797.2021.2008879>
- Min, M. (2023). School culture, self-efficacy, outcome expectation, and teacher agency toward reform with curricular autonomy in South Korea: A social cognitive approach. *Asia Pacific Journal of Education*, 43(4), 951–967. <https://doi.org/10.1080/02188791.2019.1626218>
- Mitchell, I. K., & Walinga, J. (2017). The creative imperative: The role of creativity, creative problem solving and insight as key drivers for sustainability. *Journal of Cleaner Production*, 140, 1872–1884. <https://doi.org/10.1016/j.jclepro.2016.09.162>
- Mollazehi, M., & Abdel-Salam, A.-S. G. (2024). Understanding the alternative Mantel-Haenszel statistic: Factors affecting its robustness to detect non-uniform DIF. *Communications in Statistics - Theory and Methods*, 1–25. <https://doi.org/10.1080/03610926.2024.2330668>
- Montgomery, K. (2002). Authentic tasks and rubrics: Going beyond traditional assessments in college teaching. *College Teaching*, 50(1), 34–40. <https://doi.org/10.1080/87567550209595870>
- Panayides, P., Robinson, C., & Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal*, 36(4), 611–626. <https://doi.org/10.1080/01411920903018182>
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, 15(2), 020111. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020111>
- Rahimi, R. A., & Oh, G. S. (2024). Rethinking the role of educators in the 21st century: Navigating globalization, technology, and pandemics. *Journal of Marketing Analytics*, 1–16. <https://doi.org/10.1057/s41270-024-00303-4>
- Robinson-Pant, A. (2023). Education for rural development: Forty years on. *International Journal of Educational Development*, 96, 102702. <https://doi.org/10.1016/j.ijedudev.2022.102702>

- Rusch, T., Lowry, P. B., Mair, P., & Treiblmaier, H. (2017). Breaking free from the limitations of classical test theory: Developing and measuring information systems scales using item response theory. *Information & Management*, 54(2), 189–203. <https://doi.org/10.1016/j.im.2016.06.005>
- Samson, P. L. (2015). Fostering student engagement: Creative problem-solving in small group facilitations. *Collected Essays on Learning and Teaching*, 8, 153–164. <https://doi.org/10.22329/celt.v8i0.4227>
- Selfa-Sastre, M., Pifarre, M., Cujba, A., Cutillas, L., & Falguera, E. (2022). The role of digital technologies to promote collaborative creativity in language education. *Frontiers in Psychology*, 13, 828981. <https://doi.org/10.3389/fpsyg.2022.828981>
- Soeharto, S. (2021). Development of a Diagnostic Assessment Test to Evaluate Science Misconceptions in Terms of School Grades: A Rasch Measurement Approach. *Journal of Turkish Science Education*, 18(3), 351–370. <https://doi.org/10.36681/tused.2021.78>
- Soeharto, S., & Csapó, B. (2022). Assessing Indonesian student inductive reasoning: Rasch analysis. *Thinking Skills and Creativity*, 101132. <https://doi.org/10.1016/j.tsc.2022.101132>
- Stankovic, J. A., Sturges, J. W., & Eisenberg, J. (2017). A 21st century cyber-physical systems education. *Computer*, 50(12), 82–85. <https://doi.org/10.1109/MC.2017.4451222>
- Suastra, I. W., Risiati, N. P., Adnyana, P. P. B., & Kanca, N. (2019). The effectiveness of Problem Based Learning-Physics module with authentic assessment for enhancing senior high school students' physics problem solving ability and critical thinking ability. *Journal of Physics: Conference Series*, 1171(1), 012027. <https://doi.org/10.1088/1742-6596/1171/1/012027>
- Suherman, S., & Vidákovich, T. (2022). Assessment of Mathematical Creative Thinking: A Systematic Review. *Thinking Skills and Creativity*, 101019. <https://doi.org/10.1016/j.tsc.2022.101019>
- Suherman, S., & Vidákovich, T. (2024). Relationship between ethnic identity, attitude, and mathematical creative thinking among secondary school students. *Thinking Skills and Creativity*, 51, 101448. <https://doi.org/10.1016/j.tsc.2023.101448>
- Suryanto, H., Degeng, I. N. S., Djatmika, E. T., & Kuswandi, D. (2021). The effect of creative problem solving with the intervention social skills on the performance of creative tasks. *Creativity Studies*, 14(2), 323–335. <https://doi.org/10.3846/cs.2021.12364>
- Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 1. The “most likely” measures coming from the model. *Disability and Rehabilitation*, 1–13. <https://doi.org/10.1080/09638288.2023.2169771>
- Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., & Mourey, F. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence*, 11(3), 54. <https://doi.org/10.3390/jintelligence11030054>
- Treffinger, D. J. (2007). Creative Problem Solving (CPS): Powerful Tools for Managing Change and Developing Talent. *Gifted and Talented International*, 22(2), 8–18. <https://doi.org/10.1080/15332276.2007.11673491>

- Utami, Y. P., & Suswanto, B. (2022). The Educational Curriculum Reform in Indonesia: Supporting “Independent Learning Independent Campus (MBKM)”. *SHS Web of Conferences*, 149. <https://doi.org/10.1051/shsconf/202214901041>
- Val, E., Gonzalez, I., Lauroba, N., & Beitia, A. (2019). How can Design Thinking promote entrepreneurship in young people? *The Design Journal*, 22(sup1), 111–121. <https://doi.org/10.1080/14606925.2019.1595853>
- van Hooijdonk, M., Mainhard, T., Kroesbergen, E. H., & van Tartwijk, J. (2020). Creative problem solving in primary education: Exploring the role of fact finding, problem finding, and solution finding across tasks. *Thinking Skills and Creativity*, 37, 100665. <https://doi.org/10.1016/j.tsc.2020.100665>
- Van Hooijdonk, M., Mainhard, T., Kroesbergen, E. H., & Van Tartwijk, J. (2023). Creative problem solving in primary school students. *Learning and Instruction*, 88, 101823. <https://doi.org/10.1016/j.learninstruc.2023.101823>
- Wang, M., Yu, R., & Hu, J. (2023). The relationship between social media-related factors and student collaborative problem-solving achievement: An HLM analysis of 37 countries. *Education and Information Technologies*, 1–19. <https://doi.org/10.1007/s10639-023-11763-z>
- Welter, V. D. E., Dawborn-Gundlach, M., Großmann, L., & Krell, M. (2024). Adapting a self-efficacy scale to the task of teaching scientific reasoning: Collecting evidence for its psychometric quality using Rasch measurement. *Frontiers in Psychology*, 15, 1339615. <https://doi.org/10.3389/fpsyg.2024.1339615>
- Wheeler, S., Waite, S. J., & Bromfield, C. (2002). Promoting creative thinking through the use of ICT. *Journal of Computer Assisted Learning*, 18(3), 367–378. <https://doi.org/10.1046/j.0266-4909.2002.00247.x>
- Wolcott, M. D., McLaughlin, J. E., Hubbard, D. K., Rider, T. R., & Umstead, K. (2021). Twelve tips to stimulate creative problem-solving with design thinking. *Medical Teacher*, 43(5), 501–508. <https://doi.org/10.1080/0142159X.2020.1807483>
- Yu, Y., & Duchin, F. (2024). Building a Curriculum to Foster Global Competence and Promote the Public Interest: Social Entrepreneurship and Digital Skills for American Community College Students. *Community College Journal of Research and Practice*, 48(3), 164–174. <https://doi.org/10.1080/10668926.2022.2064374>
- Zhang, L., Carter, R. A., Bloom, L., Kennett, D. W., Hoekstra, N. J., Goldman, S. R., & Rujimora, J. (2024). Are Pre-Service Educators Prepared to Implement Personalized Learning?: An Alignment Analysis of Educator Preparation Standards. *Journal of Teacher Education*, 75(2), 219–235. <https://doi.org/10.1177/00224871231201367>
- Zheng, J., Cheung, K., & Sit, P. (2024). The effects of perceptions toward Interpersonal relationships on collaborative problem-solving competence: Comparing four ethnic Chinese communities assessed in PISA 2015. *The Asia-Pacific Education Researcher*, 33(2), 481–493. <https://doi.org/10.1007/s40299-023-00744-y>
- Zieky, M. (2012). Practical questions in the use of DIF statistics in test development. In *Differential item functioning* (pp. 337–347). Routledge.